Methods for Inverse Problems:
II. Nonlinear Problems and Tikhonov regularization

Barbara Kaltenbacher, University of Klagenfurt, Austria

overview

1. Nonlinear setting and Tikhonov
2. Convergence
3. Conditions on F
4. Convergence rates

Nonlinear setting

We want to solve the operator equation

$$F(x) = y \tag{1}$$

with $F : X \to Y$, X, Y Hilbert spaces,
given noisy data $y^\delta \in Y$ s.t. $\|y^\delta - y\| \leq \delta$.
Assume that for exact data y exact solution $x^\dagger$ exists and is unique.

Discuss some aspects of methods for nonlinear problems by means of the best investigated one:
Tikhonov regularization: x_α^δ minimizer of

$$J_\alpha(x) = \|F(x) - y^\delta\|^2 + \alpha\|x - x_0\|^2 = \min_{x \in D(F)}! \tag{2}$$

x_0 ... initial guess of $x^\dagger$, J_α ... Tikhonov functional

Well-definedness and stability

F is weakly closed $:\Leftrightarrow$

$$\big((\psi_n)_{n\in\mathbb{N}} \subset \mathcal{D}(F) \;\wedge\; \psi_n \rightharpoonup \psi \;\wedge\; F(\psi_n) \rightharpoonup f\big) \\ \Longrightarrow \quad \psi \in \mathcal{D}(F) \;\wedge\; F(\psi) = f\,. \tag{3}$$

Theorem

Let $\alpha > 0$ and assume that F is weakly closed (3) and continuous. Then the Tikhonov functional (2) has a global minimizer.

Theorem

Let $\alpha > 0$ and assume that F is weakly closed (3) and continuous. For any sequence $y^k \to y^\delta$ as $k \to \infty$ the corresponding minimizers x_α^k of (2) (with y^k in place of y^δ) converge to x_α^δ.

Convergence

Theorem

Assume that F is weakly closed (3) and continuous. Let $\alpha = \overline{\alpha}(\delta)$ be chosen such that

$$\overline{\alpha}(\delta) \to 0 \quad \textit{and} \quad \delta^2/\overline{\alpha}(\delta) \to 0 \qquad \textit{as } \delta \to 0. \tag{4}$$

If y^{δ_k} is some sequence in Y such that $\|y^{\delta_k} - y\| \leq \delta_k$ and $\delta_k \to 0$ as $k \to \infty$, and if $x_{\alpha_k}^{\delta_k}$ denotes a solution to (2) with $y^\delta = y^{\delta_k}$ and $\alpha = \alpha_k = \overline{\alpha}(\delta_k)$, then $\|x_{\alpha_k}^{\delta_k} - x^\dagger\| \to 0$ as $k \to \infty$.

The same result holds for the discrepancy principle
$\underline{\tau}\delta \leq \|F(x_\alpha^\delta) - y^\delta\| \leq \overline{\tau}\delta$ with fixed $1 < \underline{\tau} < \overline{\tau}$
in place of the a priori choice (4)

Conditions on F and convexity of the Tikhonov functional

Lemma

Let the weak nonlinearity *condition [Chavent&Kunisch 1996]*

$$\forall x \in D(F)\,\forall w : x + w \in D(F) : \quad \|F''(x)[w,w]\| \leq \frac{1}{R}\|F'(x)w\|^2$$

for some $R > \delta$ hold. Then for all $\alpha > 0$, the Tikhonov functional is convex in a sufficiently small neighborhood of $x^\dagger$.

Proof:

$$J_\alpha'(x)w = 2\langle F(x) - y^\delta, F'(x)w\rangle + 2\alpha\langle x - x_0, w\rangle$$

$$J_\alpha''(x)[w,w] = 2\|F'(x)w\|^2 + 2\langle F(x) - y^\delta, F''(x)[w,w]\rangle + 2\alpha\|w\|^2$$

Lemma

Let the following Taylor remainder estimate $\forall \tilde{x}, x \in D(F)$

$$\|F(\tilde{x})-F(x)-F'(x)(\tilde{x}-x)\| \leq \min\{\frac{1}{R}\|F(\tilde{x})-F(x)\|^2,\ c\|F(\tilde{x})-F(x)\|\}$$

for some $R > 2\delta$, $c < 1 - \frac{2\delta}{R}$ *hold. Then for all* $\alpha > 0$, *the Tikhonov functional is convex in a sufficiently small neighborhood of* $x^\dagger$.

Proof:

$$\begin{aligned} &\frac{1}{2}\langle J_\alpha'(\tilde{x}) - J_\alpha'(x), (\tilde{x}-x)\rangle \\ &\quad = \quad \langle F(\tilde{x}) - F(x), F'(x)(\tilde{x}-x)\rangle + \langle F(\tilde{x}) - y^\delta, (F'(\tilde{x}) - F'(x))(\tilde{x}-x)\rangle \\ &\quad \geq \quad \|F(\tilde{x}) - F(x)\|^2 - c\|F(\tilde{x}) - F(x)\|^2 - \frac{2}{R}\|F(x) - y^\delta\|\,\|F(\tilde{x}) - F(x)\|^2 \end{aligned}$$

since $(F'(\tilde{x}) - F'(x))(\tilde{x}-x) =$
$F(\tilde{x}) - F(x) - F'(x)(\tilde{x}-x) + F(x) - F(\tilde{x}) - F'(\tilde{x})(x-\tilde{x})$.

Convergence rates and source conditions

Theorem

Assume that F' is Lipschitz

$$\|F'[x] - F'[\tilde{x}]\| \leq L\|x - \tilde{x}\| \tag{5}$$

for all $x, \tilde{x} \in D(F)$. Moreover, assume that the source condition

$$x^\dagger - x_0 = F'(x^\dagger)^* w, \ \text{with } \|w\| < \frac{1}{L}$$

is satisfied for some $w \in Y$ and that α is chosen according to
(a) the a priori parameter choice rule $\alpha = c\delta$ with some $c > 0$
(b) the discrepancy principle $\underline{\tau}\delta \leq \|F(x_\alpha^\delta) - y^\delta\| \leq \overline{\tau}\delta$ with fixed $1 < \underline{\tau} < \overline{\tau}$.
Then there exists a constant $C > 0$ independent of δ such that

$$\|x_\alpha^\delta - x^\dagger\| \leq C\sqrt{\delta}, \quad \|F(x_\alpha^\delta) - y\| \leq C\delta.$$

Theorem

Assume that F satisfies the tangential cone condition [Scherzer 1995]

$$\|F(\tilde{x}) - F(x) - F'(x)(\tilde{x} - x)\| \le c\|F(\tilde{x}) - F(x)\| \qquad (6)$$

for all $x, \tilde{x} \in D(F)$. Moreover, assume that the source condition with $\mu \in (0, \frac{1}{2})$

$$x^\dagger - x_0 = (F'(x^\dagger)^* F'(x^\dagger))^\mu w$$

is satisfied for some $w \in X$ and that α is chosen according to (a) or (b).
Then there exists a constant $C > 0$ independent of δ such that

$$\|x_\alpha^\delta - x^\dagger\| \le C\delta^{\frac{2\mu}{2\mu+1}}, \quad \|F(x_\alpha^\delta) - y\| \le C\delta.$$

An example

Estimate the diffusion coefficient a in

$$\begin{aligned} -(a(s)u'(s))' = f(s)\,, \qquad & s \in (0,1)\,, \\ u(0) = 0 = u(1)\,, & \end{aligned} \tag{7}$$

where $f \in L^2$.

$$F : \mathcal{D}(F) := \{a \in H^1[0,1] \,:\, a(s) \geq \underline{a} > 0\} \to L^2[0,1]$$

$$\begin{aligned} F(a) &= u(a) = A(a)^{-1} f\,, \\ F'(a)h &= A(a)^{-1}[(hu'(a))']\,, \\ F'(a)^* w &= -B^{-1}[u'(a)(A(a)^{-1} w)']\,, \end{aligned}$$

- $A(a) : H^2[0,1] \cap H_0^1[0,1] \to L^2[0,1]$,
  $A(a)u := -(au')'$
- B^{-1} ... adjoint of the embedding $H^1[0,1] \to L^2[0,1]$.
  $B : \{\psi \in H^2[0,1] \,:\, \psi'(0) = \psi'(1) = 0\} \to L^2[0,1]$,
  $B\psi := -\psi'' + \psi$.

$$
\begin{aligned}
\langle\, F(\tilde a) - F(a) - F'(a)(\tilde a - a), w \,\rangle_{L^2} \\
&= \langle\, (\tilde u - u) - A(a)^{-1}[((\tilde a - a)u')'], w \,\rangle_{L^2} \\
&= \langle\, A(a)(\tilde u - u) - ((\tilde a - a)u')', A(a)^{-1} w \,\rangle_{L^2} \\
&= \langle\, ((\tilde a - a)(\tilde u' - u'))', A(a)^{-1} w \,\rangle_{L^2} \\
&= -\langle\, (\tilde a - a)(\tilde u - u)', (A(a)^{-1} w)' \,\rangle_{L^2} \\
&= \langle\, F(\tilde a) - F(a), ((\tilde a - a)(A(a)^{-1} w)')' \,\rangle_{L^2} \,.
\end{aligned}
$$

$$
\begin{aligned}
&\|((\tilde a - a)(A(a)^{-1} w)')'\|_{L^2} = \|(\tfrac{\tilde a - a)}{a}\, a(A(a)^{-1} w)')'\|_{L^2} \\
&= \|(\tfrac{\tilde a - a}{a})'\, a(A(a)^{-1} w)' + \tfrac{\tilde a - a}{a}(a(A(a)^{-1} w')'\|_{L^2} \\
&\le \tfrac{1}{\underline{a}}\Big(\|(\tilde a - a)'\|_{L^2} + \tfrac{1}{\underline{a}^2}\|a'\|_{L^2}\|\tilde a - a)\|_{L^\infty} \Big) \underbrace{\|a(A(a)^{-1} w)'\|_{L^\infty}}_{\le \bar a \hat C \|(A(a)^{-1} w)'\|_{H^1} \le \bar a \hat C \|A(a)^{-1}\|_{L^2 \to H^2} \|w\|_{L^2}} + \tfrac{1}{\underline{a}}\|\tilde a - a\|_{L^\infty}\|w\|_{L^2} \Big)
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow\ & \|F(\tilde a) - F(a) - F'(a)(\tilde a - a)\|_{L^2} \\
&\le\ C\,\|F(\tilde a) - F(a)\|_{L^2}\,\|\tilde a - a\|_{H^1} \,.
\end{aligned}
$$

Literature:

- stability and convergence: [Seidman&Vogel 1989]
- convergence rates [Engl&Kunisch&Neubauer 1989] [Neubauer 1999] [Hofmann&Scherzer 1998]
- analysis in Banach space: [Burger&Osher 2004], [Hofmann&Kaltenbacher&Pöschl&Scherzer 2007] [Hein&Hofmann&Kindermann&Neubauer&Tautenhahn 2009]
- ...